JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 51, 440–448 (1975)

Approximation Numbers of Bounded Operators

EBERHARD SCHOCK

Fachbereich Mathematik der Universität, D 6750 Kaiserslautern, West Germany

Submitted by Ky Fan

DEDICATED TO PROFESSOR H. UNGER ON THE OCCASION OF HIS 60TH BIRTHDAY

The theory of ideals of linear operators is well developed and has a lot of applications in theory and practise. The purpose of this paper is to give a first idea of a similar theory for bounded (nonlinear) operators. In view of applications we will not give an abstract (perhaps general nonsense) theory, but an example of a class λ_p of bounded operators with a structure similar to an $\mathscr{L}$-module ($\mathscr{L}$ represents the class of all linear operators between Banach spaces), and applications to projection methods for solving equations with λ_p-type operators.

1.

Let M be a subset of a Banach space X, and Y a Banach space. By $B(M, Y)$ we denote the Banach space of all bounded continuous operators $T: M \to Y$ with the norm $s(T) = \sup\{\| Tx \|, x \in M\}$, and by $B_n(M, Y)$ the set of all $T \in B(M, Y)$ with the property: TM is contained in an n-dimensional subspace of Y. $\mathscr{L}(X, Y)$ is the usual Banach space of all continuous linear operators $L: X \to Y$. Then we define for $T \in B(M, Y)$

$$a_n(T) = \inf\{s(T - T_n), T_n \in B_n(M, Y)\},$$

the nth *approximation number* of T. Evidently,

$$s(T) \geqslant a_0(T) \geqslant a_1(T) \geqslant a_2(T) \geqslant \cdots \geqslant 0.$$

Approximation numbers of linear operators were introduced by Pietsch (7):

$$\alpha_n(L) = \inf\{\| L - L_n \|, L_n \in \mathscr{L}(X, Y), \dim L_n X \leqslant n\}$$

for $L \in \mathscr{L}(X, Y)$.

The following properties are evident.

(A1) For all $S, T \in B(M, Y)$, for all $n, m \in N$,

$$a_{n+m}(S + T) \leqslant a_n(S) + a_m(T).$$

(A2) For all S, $T \in B(M, Y)$, for all $n \in N$,

$$|a_n(S) - a_n(T)| \leqslant s(S - T).$$

(A3) For all $T \in B(M, Y)$, for all scalars c,

$$a_n(cT) = |c| \, a_n(T).$$

(A4) For all $T \in B(M, Y)$, for all $n \in N$,

$$a_n(T) = 0 \qquad \text{iff} \quad T \in B_n(M, Y).$$

(A5) For all $T \in B(M, Y)$, for all $L \in \mathscr{L}(Y, Z)$, $n \in N$,

$$a_n(LT) \leqslant \|L\| \, a_n(T).$$

PROPOSITION 1.1. *Let b: $B(M, Y) \to R^N$ be a mapping with the properties* (A1) *and* (A4). *Then for all $n \in N$ and $T \in B(M, Y)$,*

$$b_n(T) \leqslant a_n(T).$$

Proof. (A1) implies with $S \in B(M, Y)$

$$b_n(T) \leqslant b_n(S) + b_0(S - T) \leqslant b_n(S) + s(S - T).$$

For $\epsilon > 0$ choose $S \in B_n(M, Y)$ with $s(S - T) \leqslant a_n(T) + \epsilon$, then

$$b_n(T) \leqslant s(S - T) \leqslant a_n(T) + \epsilon,$$

hence the mapping a: $B(M, Y) \to R^N$ with $a(T) = (a_n(T))$ is the greatest approximation number function.

Let A be a bounded set in a Banach space Y. By $\chi(A)$, the *measure of noncompactness of* A (3), we denote the infimum of all $\epsilon > 0$, such there exists a finite ϵ-net for A.

PROPOSITION 1.2. *For $T \in B(M, Y)$,*

$$\chi(TM) = \lim_{n \to \infty} a_n(T).$$

Proof. Let $\chi := \lim_{n\to\infty} a_n(T)$ and U the unit ball in Y.

(a) "$\chi(TM) \geqslant \chi$": For $\delta > 0$ there exists a finite set $y_1, \ldots, y_k$ with

$$TM \subset \bigcup_{j=1}^{k} (y_j + (\chi(TM) + \delta)\, U).$$

We define $\lambda_j(x) = \max(0, \delta - \| Tx - y_j \|)$ and

$$\tau_j(x) = \left(\sum_{i=1}^{\infty} \lambda_i(x)\right)^{-1} \lambda_j(x).$$

Then we have for $T_k(x) = \sum_{j=1}^{k} \tau_j(x)\, y_j$, $T_k \in B_k(M, Y)$ and $s(T - T_k) \leqslant \chi(TM) + \delta$. Hence $a_k(T) \leqslant \chi(TM) + \delta$, and therefore $\chi \leqslant \chi(TM)$.

(b) "$\chi(TM) \leqslant \chi$": $\chi = \lim_{n\to\infty} a_n(T)$ implies: for $\delta > 0$ there exists $n_0 \in N$ such that for $n \geqslant n_0$, $a_n(T) - \chi < \delta/2$. Hence there exists $T_n \in B_n(M, Y)$ with $s(T - T_n) - \chi < \delta/2$. T_nM is relatively compact in Y; therefore there is a finite set $y_1, \ldots, y_k \in Y$ with

$$T_nM \subset \bigcup_{j=1}^{k} y_j + \frac{\delta}{2}\, U,$$

hence

$$TM \subset \bigcup_{j=1}^{k} y_j + \frac{\delta}{2}\, U + \left(\chi + \frac{\delta}{2}\, U\right) = \bigcup_{j=1}^{k} y_j + (\chi + \delta)\, U.$$

So we obtain a finite $(\chi + \delta)$-net for TM, hence $\chi(TM) \leqslant \chi$.

COROLLARY. *T is compact if and only if* $\lim_{n\to\infty} a_n(T) = 0$.

In the sequel we will use the following composition formulas.

PROPOSITION 1.3. *Let $T \in B(M, Y)$, $S \in B(N, Z)$ with $TM \subset N$, Z a Banach space. Let S be Lipschitz continuous*: $\| Sy' - Sy'' \| \leqslant k \cdot \| y' - y'' \|$. *Then $a_n(ST) \leqslant k \cdot a_n(T)$.*

Proof. For $\epsilon > 0$ choose $T_n \in B_n(M, Y)$ with $s(T - T_n) \leqslant a_n(T) + \epsilon$. Then $a_n(ST) \leqslant s(ST - ST_n) \leqslant k \cdot s(T - T_n) \leqslant k(a_n(T) + \epsilon)$.

PROPOSITION 1.4. *Let X, Y, Z be Banach spaces, $M \subset X$, $T \in B(M, Y)$ and $L \in \mathscr{L}(Y, Z)$. Then for $n, m \in N$*

$$a_{n+m}(LT) \leqslant \alpha_n(L)\, a_m(T).$$

Proof. For $\epsilon > 0$ choose operators $L_n \in \mathscr{L}(Y, Z)$ and $T_m \in B_m(M, Y)$ with $\dim L_nY \leqslant n$, $s(T - T_m) \leqslant a_m(T) + \epsilon$, and $\| L - L_n \| \leqslant \alpha_n(L) + \epsilon$. Then $L_n(T - T_m) + LT_m \in B_{n+m}(M, Z)$ and

$$\begin{aligned} a_{n+m}(LT) &\leqslant s(LT - L_n(T - T_m) - LT_m) \\ &\leqslant s((L - L_n)(T - T_m)) \leqslant \| L - L_n \| \cdot s(T - T_m) \\ &\leqslant (\alpha_n(L) + \epsilon)(a_m(T) + \epsilon). \end{aligned}$$

Pietsch (7) introduced the following class of linear operators. A linear operator L is said to be of type l^p if and only if $\sum \alpha_n(L)^p < \infty$ $(p > 0)$. The set $l^p(X, Y)$ of all operators $L \in \mathscr{L}(X, Y)$ of type l^p is a linear quasinormed (by $\rho_p(L) = (\sum \alpha_n(L)^p)^{1/p}$) complete space. Analogs we define: $T \in B(M, Y)$ is said to be of *type* λ_p, if and only if

$$\sigma_p(T) := \left(\sum a_n(T)^p\right)^{1/p} < \infty.$$

We denote the set of all λ_p-type operators by $\lambda_p(M, Y)$.

PROPOSITION 1.5. *$\lambda_p(M, Y)$ is a linear space, σ_p is a quasinorm on $\lambda_p(M, Y)$, and $\lambda_p(M, Y)$ is, with Y, complete in the topology, induced by σ_p.*

The proof is the same as in (7, Chap. 8.2). Let us remark, that $\lambda_p(M, Y)$ has an "$\mathscr{L}$-module property" ($L \in \mathscr{L}$ and $T \in \lambda_p$ implies $LT \in \lambda_p$).

PROPOSITION 1.6. *Let $L \in L(Y, Z)$, $T \in \lambda_p(M, Y)$, then $LT \in \lambda_p(M, Z)$, and*

$$\sigma_p(LT) \leqslant \| L \| \, \sigma_p(T).$$

The proof follows from (A5) and from the definition of $\lambda_p(M, Y)$.

The next proposition shows, that the class of λ_p-type operators is not empty.

PROPOSITION 1.7. *Let $L \in l^p(Y, Z)$ and $T \in B(M, Y)$. Then $LT \in \lambda_p(M, Z)$ and*

$$\sigma_p(LT) \leqslant s(T) \, \rho_p(L).$$

Proof. From Proposition 1.4 follows $a_n(LT) \leqslant \alpha_n(L) \cdot s(T)$, hence $\sigma_p(LT) = (\sum a_n(LT)^p)^{1/p} \leqslant s(T)(\sum \alpha_n(L)^p)^{1/p} \leqslant s(T) \, \rho_p(L) < \infty$.

PROPOSITION 1.8. *Let $L \in l^p(Y, Z)$ and $T \in \lambda_q(M, Y)$. Then $LT \in \lambda_r(M, Z)$ with $1/r = 1/p + 1/q$, and $\sigma_r(LT) \leqslant \tau_{pq} \rho_p(L) \cdot \sigma_q(T)$, where τ_{pq} is independent from L and T.*

The proof is given in (8).

The next proposition shows, that, if $0 < p \leqslant 1$, each λ_p-type operator can be expanded in an absolute convergent series.

PROPOSITION 1.9. *Let $T \in \lambda_p(M, Y)$, $0 < p \leqslant 1$. Then there exist a* *sequence* (f_n) *of scalar-valued continuous-bounded functions, and a sequence* $(y_n) \subset Y$ *with* $\| y_n \| = 1$, *such that* T *has the form*

$$Tx = \sum_{n=1}^{\infty} f_n(x)\, y_n \qquad \text{for } x \in M$$

with

$$\left(\sum_{n=1}^{\infty} \| f_n \|^p\right)^{1/p} \leqslant 2^{2+3/p}\sigma_p(T).$$

The proof is also given in (8).

2.

In this section we describe methods to estimate the approximation numbers of concrete operators.

PROPOSITION 2.1. *Let* $T \in B(M, Y)$. *If there exists a linear operator* D *in* Y *with continuous inverse and a positive number* m, *such that for all* $x \in M$

$$\| DTx \| \leqslant m,$$

then

$$a_n(T) \leqslant m(1 + n^{1/2})\, \alpha_n(D^{-1}).$$

Proof. Let $U = \{y \in Y, \| y \| \leqslant 1\}$ and $U_D = \{y \in Y, \| Dy \| \leqslant 1\}$. For $\epsilon > 0$ and $n \in N$ we choose $L_n \in \mathscr{L}(Y, Y)$ with $\dim L_n Y \leqslant n$, and $\| D^{-1} - L_n \| \leqslant (1 + \epsilon)\, \alpha_n(D^{-1})$. Since $y \in U$ implies $D^{-1}y \in U_D$, we have

$$U_D \subset \| D^{-1} - L_n \| \cdot U + L_n Y.$$

By assumption $TM \subset mU_D$, so we have with $Y_n = L_n Y$

$$TM \subset m(1 + \epsilon)\, \alpha_n(D^{-1}) \cdot U + Y_n .$$

By a theorem of Kadec and Snobar (4) there exists a linear projection $P_n \colon Y \to Y_n$ with $\| P_n \| \leqslant n^{1/2}$. For $x \in M$ we get

$$Tx = m(1 + \epsilon)\, \alpha_n(D^{-1}) \cdot v + w$$

with $v \in U$ and $w \in Y_n$; this implies with $P_n w = w$

$$\begin{aligned} \| Tx - P_n Tx \| &= m(1 + \epsilon)\, \alpha_n(D^{-1}) \cdot \| v - P_n v \| \\ &\leqslant m(1 + \epsilon)\, \alpha_n(D^{-1})(1 + n^{1/2})\, \| v \|, \end{aligned}$$

and we obtain

$$a_n(T) \leqslant s(T - P_n T) \leqslant m(1 + n^{1/2})\, \alpha_n(D^{-1}).$$

Corollary. *If X is strictly convex or reflexive, then*

$$\| DTx \| \leqslant m \text{ implies } a_n(T) = O(\alpha_n(D^{-1})).$$

Proof. If X is reflexive, choose an equivalent norm on X which is strictly convex. Then choose P_n to be the (homogeneous) metric projection (10). Let Ω be an open bounded domain in m-dimensional Euclidean space with sufficiently smooth boundary. Then for $k > 0$ and $1 \leqslant p < \infty$ the Sobolev space $W_p^k(\Omega)$ is the completion of the space $C^\infty(\bar{\Omega})$ with respect to the norms

$$\| x \|_{W_p^k} = \sum_{|\alpha| \leqslant k} \| D^\alpha x \|_{L_p} \qquad \text{if} \qquad k \in N,$$

and

$$\| x \|_{W_p^k} = \| x \|_{W_p^{[k]}} + \left(\sum_{|\alpha| = [k]} \int_\Omega \int_\Omega \frac{| D^\alpha x(t) - D^\alpha x(\tau)|^p}{| t - \tau |^{m + (k - [k])p}}\, dt\, d\tau \right)$$

if $k \notin N$.

Proposition 2.2. *Let $\hat{T}$ be a continuous-bounded operator from a subset M of $L_p(\Omega)$ into $W_p^k(\Omega)$, κ the embedding of $W_p^k(\Omega)$ in $L_p(\Omega)$ and $T = \kappa \hat{T}$. Then,*

$$a_n(T) = O(n^{-k/m}).$$

Proof. We have the following commutative diagram.

$$\begin{tikzcd}
M \subset L_p(\Omega) \arrow[rr, "T"] \arrow[dr, "\hat{T}"'] & & L_p(\Omega) \\
& W_p^k(\Omega) \arrow[ur, "\kappa"'] &
\end{tikzcd}$$

By a theorem of Birman and Solomjak (2) we have

$$\alpha_n(\kappa) = O(n^{-k/m}),$$

and therefore, by Proposition 1.4,

$$a_n(T) \leqslant \alpha_n(\kappa)\, s(T) = O(n^{-k/m}).$$

PROPOSITION 2.3. *Let $\hat{T}$ be a continuous bounded operator on a bounded subset M of $L_p(\Omega)$ into $W_p{}^k(\Omega)$. Suppose, that $\hat{T}$ fulfills a Lipschitz condition on M with Lipschitz constant c. Let T be the restriction of $\hat{T}$ to $M \cap W_p{}^k(\Omega)$. Then,*

$$a_n(T) = O(n^{-k/m}).$$

Proof. We have

$$\begin{tikzcd} M \cap W_p{}^k(\Omega) \arrow[rr, "T"] \arrow[dr, "\kappa"'] & & W_p{}^k(\Omega) \\ & L_p(\Omega) \arrow[ur, "\hat{T}"'] & \end{tikzcd}$$

where κ is the embedding of $M \cap W_p{}^k(\Omega)$ into $L_p(\Omega)$. Then $T = \hat{T}\kappa$. By the theorem of Birman and Solomiak there exists a linear operator $\kappa_n\colon W_p{}^k(\Omega) \to L_p(\Omega)$ with $\| \kappa - \kappa_n \| = O(n^{-k/m})$; therefore we have

$$\begin{aligned} a_n(T) &\leqslant s(T\kappa - T\kappa_n) \leqslant c \sup\{\| \kappa x - \kappa_n x \|, x \in M\} \\ &\leqslant c \| \kappa - \kappa_n \| \sup\{\| x \|, x \in M\} = O(n^{-k/m}). \end{aligned}$$

PROPOSITION 2.4. *Let T be a continuous-bounded operator from a subset M of $W_p{}^k(\Omega)$ into $W_p{}^j(\Omega)$ $(j > k)$, κ the embedding of $W_p{}^j(\Omega)$ into $W_p{}^k(\Omega)$, and $T = \kappa\hat{T}$. Then,*

$$a_n(T) = O(n^{-(j-k)/m+1/2}).$$

Proof. By a theorem of Amar el Kolli (5) the nth diameter of the unit ball $U_p{}^j$ of $W_p{}^j(\Omega)$ with respect to the $W_p{}^k$-norm is $d_n(U_p{}^j, U_p{}^k) = O(n^{-(j-k)/m})$, where $d_n(A, B)$ is given by $d_n(A, B) = \inf\{\inf\{t > 0, A \subset tB + F\}, F \subset W_p{}^k, \dim F \leqslant n\}$. Lemma 9.1.6 of (7) states the result:

$$\alpha_n(\kappa) \leqslant (1 + n^{1/2})\, d_n(U_p{}^j, U_p{}^k) = O(n^{-(j-k)/m+1/2}),$$

therefore follows $a_n(T) = O(n^{-(j-k)/m+1/2})$.

Similar to Proposition 2.3 is the following.

PROPOSITION 2.5. *Let $\hat{T}$ be a continuous bounded operator from a bounded subset M of $W_p{}^k(\Omega)$ into $W_p{}^j(\Omega)$ $(j > k)$. Suppose that $\hat{T}$ fulfills a Lipschitz condition on M. Let T be the restriction of $\hat{T}$ to $M \cap W_p{}^j(\Omega)$. Then*

$$a_n(T) = O(n^{-(j-k)/m+1/2}).$$

3.

As an application we will show that the optimal rate of convergence of some projection method for solving an equation

$$x - Tx = y \tag{\&}$$

is given by the rate of convergence of the approximation numbers of T.

For simplicity we assume, that $T \in B(K_s, Y)$, with $K_s = \{x \in Y \mid \| x \| \leqslant s\}$, is Lipschitz continuous with Lipschitz constant $k < 1$. Then, by the contraction principle, for all $y \in K_r$ with $r = (1 - k)s$, the equation (&) has a unique solution $x^\dagger$. The mapping $R: K_r \to K_s$ with $Ry = x^\dagger$ is called the resolvent of T (6). It is known: $R \in B(K_r, Y)$, and for all $y \in K_r$,

$$y = Ry - TRy.$$

By $Sy := Ry - y$ for $y \in K_r$ we define the quasiresolvent $S \in B(K_r, Y)$ of T. Then we have

$$S = TR.$$

This means especially: if T is compact, then S is also compact. Now we show $a_n(T) \geqslant a_n(S)$. For $\epsilon > 0$ there is $T_n \in B_n(K_s, Y)$ with $s(T - T_n) \leqslant a_n(T) + \epsilon$. Then

$$\begin{aligned} s(S - T_nR) &= \sup\{\| Sy - T_nRy \|, y \in K_r\} \\ &= \sup\{\| TRy - T_nRy \|, y \in K_r\} \\ &= \sup\{\| Tx - T_nx \|, x \in K_s\} \leqslant a_n(T) + \epsilon. \end{aligned}$$

On the other hand, if R is surjective, we find for $\epsilon > 0$, $S_n \in B_n(K_r, Y)$ with $s(S - S_n) \leqslant a_n(S) + \epsilon$. Then we obtain

$$\begin{aligned} s(T - S_n(I - T)) &= \sup\{\| Tx - S_n(x - Tx)\|, x \in K_s\} \\ &= \sup\{\| TRy - S_n(Ry - TRy)\|, y \in K_r\} \\ &= \sup\{\| Sy - S_ny \|, y \in K_r\} \leqslant a_n(S) + \epsilon. \end{aligned}$$

So we have $a_n(S) = a_n(T)$.

A projection method of the class Q_1 (9) consists of determining the solution $z \in Y_n = \operatorname{span}\{y_1, \ldots, y_n\}$ of the equation

$$x_j'(z - T(y + z)) = 0, \tag{\&\&}$$

where $x_1', \ldots, x_n'$ are continuous linear functionals. Then $x_n = y + z$ is an approximation of the solution x of (&). If we assume, that for each $y \in K_r$

the solution z of (&&) is uniquely determined, then a map $S_n: K_r \to Y_n$ is given. We estimate the error:

$$\| x - x_n \| = \| Ry - x_n \| = \| y + Sy - (y + S_n y) \| = \| Sy - S_n y \|.$$

So we obtain the result:

The (optimal) rate of convergence of a projection method of the class Q_1, uniformly for all $y \in K_r$, is given by the rate of convergence of the sequence $(a_n(T))$.

Remark. If we assume, that T is a Hammerstein operator $T = Lf$ with a compact self-adjoint operator L in a Hilbert space, then, by Proposition 1.4, we obtain a result similar to Amann (1).

REFERENCES

1. H. AMANN, Über die Konvergenzgeschwindigkeit des Galerkin-Verfahrens für die Hammersteinsche Gleichung, *Arch. Rational Mech. Anal.* **37** (1970), 33–47.
2. M. S. BIRMAN AND M. Z. SOLOMYAK, Piecewise polynomial approximation of functions of the class W_p^α, (russ.) *Mat. Sb.* **73** (1967), 331–335.
3. G. DARBO, Punti uniti in transformzioni a condominio non compatto, *Rend. Sem. Mat. Univ. Padova* **24** (1955), 84–92.
4. M. I. KADEC AND M. G. SNOBAR, Über einige Funktionale auf dem Minkowski-Kompaktum (russ.), *Mat. Zametki* **10** (1971), 453–458.
5. AMAR EL KOLLI, n^{me}, épaisseur dans les espaces de Sobolev, *C. R. Acad. Sci. Paris* **272** (1971), 537–539.
6. M. A. KRASNOSELSKI, "Topological Methods in the Theory of Nonlinear Integral equations," Pergamon, New York, 1963.
7. A. PIETSCH, "Nukleare lokalkonvexe Räume," Akademie Verlag, Berlin, 1965.
8. E. SCHOCK, "Über einige lineare Räume von nichtlinearen Abbildungen," Westdeutscher Verlag, Köln–Opladen, 1967.
9. E. SCHOCK, On projection methods for linear equations of the second kind, *J. Math. Anal. Appl.* **45** (1974), 293–299.
10. I. SINGER, "Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces," Springer-Verlag, Berlin, 1970.